Abstract

Methods are given for improving discriminative training of hidden Markov models for continuous speech recognition. In one approach, discriminatively trained mixture models are interpolated with maximum likelihood trained mixture models. In another approach, segmentation and recognition results from one set of models are reused to discriminatively train a second set of models. For example, segmentation and recognition results from detailed match models are mapped and used to discriminatively train fast match models. In addition, gradients for the standard deviation of mixture components are clipped based on the statistics of the gradients. Pronunciation of words may also be used to determine the "incorrect" recognition hypothesis.